REMARKS

The Office Action dated August 13, 2001, has been carefully considered. Accordingly, the changes presented herewith, taken with the following remarks, are believed sufficient to place the present Application in condition for allowance. Reconsideration is respectfully requested.

Claims 1-7, 10, and 22-31 are pending in the case. Upon entry of this Amendment, Claim 10 will be amended, and Claims 34-36 will be newly added. Support for the amendment to Claim 10 can be found in original Claim 3 among other places. Support for Claims 34 and 36 can be found in original Claims 3 and 10 among other places. Support for Claim 35 may be found in original Claim 1, among other places. As the amendments to the claims are supported by the specification as originally filed and add no new matter, entry is believed to be in order.

I. The First 35 U.S.C. Section 112 Rejection in the Office Action

Beginning with Page 3 and continuing through Page 6, the Office Action rejects Claims 1-10 and 22-32 under 35 U.S.C. § 112, first paragraph, as containing subject matter not described in the specification.

More particularly, beginning with Page 3 and continuing to Page 5, the Office Action raises an issue with respect to whether the specification is sufficiently enabling for a hybridization clause in the claims. Then, beginning with Page 5 and continuing to Page 6, the Office Action raises an issue with respect to the use of the term "insect homo-oligomeric acetylcholine receptor", and whether those skilled in the art will know what is meant when the term "insect" is used to modify the term "homo-oligomeric acetylcholine receptor." The following discussion will address each issue in turn.

A. Hybridization Issue

Beginning with Page 3 and continuing through Page 5 of the Office Action, the Office Action alleges that the SEQ ID NOS; 1, 3 and 5 display substantial variability and would not hybridize to each other. The Examiner concludes, therefore, that the claims are drawn to three distinct genuses of nucleic acids, and that the specification has not adequately described each of these genuses. The

Examiner has indicated this portion of the rejection may be overcome by deleting the hybridization clause from the claims.

By way of response, Applicants respectfully traverse the Office Action's reasoning and its conclusion. The heart of the Office Action's reasoning is that the sequences (SEQ ID NOS 1, 3 and 5) are three distinct genuses of nucleic acids based upon sequence comparisons prepared by the Examiner which the Office Action alleges indicates a low overall sequence identity of SEQ ID NOS 1, 3 and 5 to each other.

However, Applicants wish to call to the Examiner's attention that when only partial sequences (between positions 1295 and 2195 of SEQ ID No 1, position 432 and position 1318 of SEQ ID No 3 and position 154 and position 1123 of SEQ ID N0O 5) are compared, the sequence identity is approximately 70%. See the table below. Full sequence comparisons are given in **Attachment 1** to this Amendment. The pairwise alignments were performed using the Bestfit program from the Wisconsin Package with standard parameters.

TABLE 1

COMPARISON OF PARTIAL SEQUENCES OF SEQ ID NOS 1, 3 and 5

	SEQ ID 1	SEQ ID 3	SEQ ID 5
SEQ ID 1		74	69
SEQ ID 3	74		74
SEQ ID 5	69	71	

Whereas SEQ ID NOS 1, 3 and 5 are complete cDNAs as cloned from the original cDNA libraries, the sub-sequences described above code for protein regions which are important for the function of acetylcholine receptors, i.e. the extracellular part and transmembrane regions. To those skilled in the art, this is evident from the translations into protein supplied with SEQ ID NOS 1, 3 and 5. It is well known that regions of cDNAs coding for functionally important parts of proteins are more conserved than non-coding regions.

The Office Action alleges that the sequences would not hybridize to each other under the conditions of embodiment (b) of Claim 1, again based on sequence comparisons. However, using the formula for calculating hybridization conditions originally given by Bonner et al and found in standard textbooks (for instance Sambrook et al, Molecular Cloning, 2nd Edition, CSHL Press, 1989) it becomes obvious that sequences approximately 600 pb long with approximately 70% identity and 50% G+C content, **do hybridize** in 2xSS at 60 °C. SEQ ID NOS 1, 3, AND 5 contain sub-sequences that fulfill this requirement.

The Office Action continues by stating that neither the specification nor the prior art teaches any specific correlation between any physical structure of the nucleic acid and the ability of the encoded product to form a complete or partial homo-oligomeric receptor. Applicants do not appreciate the logic of this reasoning. On pages 1 and 2 of the specification, there is referred to prior art which describes such correlations between sequence and function. Specifically it is stated:

> This classification is also supported by an examination of the relatedness of the gene sequences of the different subunits. Typically, the sequences of functionally homologous subunits from different species are more similar to each other than are sequences of subunits which are from different groups but from the same species. Thus, the rat muscle α subunit, for example, exhibits 78% amino acid identity and 84% amino acid similarity with that of the electric ray Torpedo californica but only 48% identity an 59% similarity with the rat $\alpha 2$ subunit (hetero-oligomeric, neuronal) and 36% identity and 45% similarity with the rat $\alpha 7$ subunit (homooligomeric, neuronal).

The example above compares protein sequences. Similar comparisons can be performed on nucleic acid sequences by anyone sufficiently skilled in the art using commonly accessible sequence comparison programs.

To illustrate the foregoing, attached as **Attachment 2** to this Amendment, is a comparison of coding regions performed with the program ClustalX 1.8. The sequences analyzed are SEQ ID NOS 1, 3 and 5 (insect homo-oligomeric), human and chick alpha-7 subunits (vertebrate homo-oligomeric), various insect hetero-oligomeric alpha subunits and all known hetero-oligomeric alpha subunits from

human. The output of the program contains a tree-diagram which indicates the relatedness of the sequences. It can be clearly seen from the diagram that SEQ ID NOS 1, 3 and 5 share identifying characteristics which are made evident by the sequence comparison. They furthermore fulfill the structure-function-relationship typical of the gene family of acetylcholine receptors.

In summary, in the present application it has been conclusively shown that SEQ ID NOS 1, 3 and 5 represent three species (cloned from two different evolutionary divergent insect species, one a dipteran, the other a lepidoteran) of the same genus, i.e. cDNAs encoding subunits of homo-oligomeric insect acetylcholine receptors. They share identifying characteristics which are made evident by the sequence comparison. Thus the present application provides sufficient description to indicate possession of the invention by the Applicants to the full breadth of the claims.

B. Insect Homooligomeric Acetylcholine Receptor/Specification Issue

Beginning on page 5 and continuing through page 6, the Office Action alleges that the specification fails to provide an adequate written description of what distinguishes an "insect" homooligomeric acetylcholine receptor from a "non-insect" homo-oligomeric acetylcholine receptor. The Office Action again bases much of its conclusion on an alleged high degree of variability between SEQ ID NOS 1, 3 and 5. The Office Action also bases much of its conclusion on its own prior art search of alleged mammalian homologues which the Office Action alleges, have a similar level of identity with the claimed sequences. The Office Action states that this portion of the rejection may be overcome by deleting the term "insect" from the claims.

Applicants respectfully traverse, and assert that those skilled in the art will fully understand what is meant by an insect homooligomeric acetylcholine receptor, and that the specification is not deficient in that regard and fully supports the claims. Applicants point out that the alleged high degree of variability between SEQ ID NOS 1, 3 and 5, as demonstrated above, does not exist rendering the Office Action's arguments and conclusions inapplicable. Applicants also respectfully refer the Examiner further to the discussion below in connection with the rejection of Claims 1-7, 10 and 22-31 on a similar basis.

II. The Second 35 U.S.C. Section 112 Rejection in the Office Action

Beginning on Page 6 under the heading "Enablement" and continuing through Page 8, the Office Action rejects Claims 1-10 and 22-31under 35 U.S.C. § 112, first paragraph, as not being supported by an enabling specification. (Applicants point out that Claims 1-7, 10 and 22-31 are actually pending in the case, with the addition of Claims 34-36 upon entry of this Amendment). The Office Action alleges that neither the specification nor the prior art provide guidance as to how to distinguish insect from non-insect sequences based solely on the sequence information, or disclose any characteristics unique to insect sequences. The Examiner has indicated the rejection may be overcome by deleting the term "insect" from the claims.

By way of response, Applicants respectfully traverse this ground of rejection of the claims. The Office Action alleges that one could not identify the source (insect or non-insect) of a given acetylcholine receptor protein based solely on the sequence. However, methods for sequence comparisons are described in the art which allow the examination of relatedness, which are referred to in the specification, and which are illustrated above in this Amendment and in Attachment 2 to this Amendment. These methods group insect homo-oligomeric receptors in one subgroup of homo-oligomeric receptors and the corresponding vertebrate receptors **in another subgroup.** This analysis can easily be extended to other , even as yet unknown sequences.

The Office Action also alleges that undue experimentation would be required to make the nucleic acids encoding homo-oligomeric insect acetylcholine receptors, because an experimenter would have to isolate these nucleic acids from and insect source.

Applicants respectfully state that they cannot follow the logic of the Office Action on this last point. The approach described by the Office Action as "undue experimentation" is in fact the precise type of experimentation that is followed by any and every experimenter who wants to obtain a particular nucleic acid from a given species. The approach is well known to those skilled in the art and is described in standard textbooks such as Sambrook et al (a portion of which is reproduced in

Attachment 3 of this Amendment). This approach has for instance been used to obtain cDNAs for acetylcholine receptor subunits from mammals using the corresponding Torpedo sequences as starting points.

In the present specification, Applicants have provided sufficient description so that anyone skilled in the art can obtain nucleic acids coding for homo-oligomeric acetylcholine receptors from insect species.

Example 1 provides a description of how to obtain RNA and cDNA libraries from Heliothis virescens using commonly applicable methods and reagents. These methods can easily be applied to other insect species. The experimenter can then perform a low stringency PCR-reaction using the methods and primer sequences described in Example 1. The resulting PCR product can be used to obtain the full-length cDNA clone or clones as described in Example 1.

Alternatively, the experimenter could use sub-sequences of SEQ ID Nos 1, 3 or 5 as probes to directly obtain the full-length cDNA clone or clones following the approached described in Sambrook et al. Those skilled in the art know how to identify useful sub-sequences in SEQ ID Nos 1, 3 or 5. For instance, sequences between position 1295 and position 2195 of SEQ ID No 1, position 432 and position 1318 of SEQ ID NO 3, and position 154 and position 1123 of SEQ ID NO. 5 are useful for this purpose. Once the experimenter has determined the sequence of the nucleic acids isolated as described above, sequence comparison of the nucleic acids with the sequences between position 1295 and position 2195 of SEQ ID NO 1, position 432 and position 1318 of SEQ ID NO 3 and position 154 and position 1123 of SEQ ID NO 5 will provide information on the relatedness of the new sequence to the sequences disclosed in the present application. To determine whether the new cDNAs functionally encode subunits of homo-oligomeric acetylcholine receptors, the test described in Example 2 can be applied. Briefly this consists of generating an expression construct containing the new cDNA, transferring it into suitable host cells and culturing the cells in selective media to obtain a resistant cell clone. The cells are then stimulated with acetylcholine as described in Example 2. If the cells respond to acetylcholine stimulation in a manner analogous to that described in Fig.

1, this indicates that the expressed cDNA encodes a functional homo-oligomeric acetylcholine receptor.

In summary, the present specification is fully sufficiently enabling for the claim of insect homo-oligomeric acetylcholine receptor subunit sequences.

Claim 10 has been rejected under 35 U.S.C. § 112, first paragraph, as not being supported by an enabling specification. The Examiner alleges that the specification is enabling for a process of preparing a polypeptide encoded by a nucleic acid of Claim 1 wherein the nucleic acid is operably linked to an expression control sequence, but does not reasonably provide enablement for the expression of polypeptides encoded by a nucleic acid of Claim 1 wherein the nucleic acid is not operably linked to an expression control sequence. The Examiner has indicated the rejection may be overcome by amending Claim 10 to require operable linkage to a transcription control sequence, as recited in Claim 3.

Claim 10 has been amended in accordance with the Examiner's recommendations, whereby the rejection has been overcome.

For the reasons set forth above, Applicants submit that the claims herein are adequately described and supported by an enabling specification. The Examiner is therefore requested to withdraw the rejections to the claims and to allow the application to pass to issue.

Claims 34-36 have been added to claim the present invention in more varying scope.

Attached is a marked-up version of the changes made to the claims by the current amendment.

Respectfully submitted,

By _____

Raymond J. Harmuth
Attorney for Applicants
Reg. No. 33,896

Bayer Corporation
100 Bayer Road
Pittsburgh, Pennsylvania 15205-9741
(412) 777-8366
FACSIMILE PHONE NUMBER:
(412) 777-8363
s:/sr/rjh0079

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE CLAIMS:

Please amend Claim 10 as follows. Please add new Claims 34-36 as follows.

10. (Twice Amended) A process for preparing a polypeptide encoded by a nucleic acid of Claim 1 comprising

(a) culturing a vector comprising at least one nucleic acid of Claim 1, wherein the nucleic acid is functionally linked to regulatory sequences which ensure expression of the nucleic acid in prokaryotic or eukaryotic cells ~~a host cell containing a nucleic acid of Claim 1 or a vector comprising at least one nucleic acid of Claim 1 under conditions which ensure expression of the nucleic acid of Claim 1~~, and

(b) isolating the polypeptide from the cell or ~~the~~ culture medium.

34. (New) A process for preparing a polypeptide encoded by a nucleic acid of Claim 1 comprising

(a) culturing a host cell comprising at least one nucleic acid of Claim 1, wherein the nucleic acid is functionally linked to regulatory sequences which ensure expression of the nucleic acid in prokaryotic or eukaryotic cells, and

(b) isolating the polypeptide from the cell or culture medium.

35. (New) An isolated nucleic acid comprising a sequence selected from

(a) a sequence according to nucleotide No. 372 to nucleotide No. 2681 of SEQ ID NO: 1, nucleotide No. 335 to nucleotide No. 1822 of SEQ ID NO: 3 and nucleotide No. 95 to nucleotide No. 1597 of SEQ ID NO: 5,

(b) a sequence complementary to the sequences defined under (a), and

(c) a sequence which, due to degeneracy of the genetic code, encodes the same amino acid sequences as those encoded by the sequences defined under (a),

wherein said nucleic acid encodes a complete or partial acetylcholine receptor subunit having the ability to form homooligomeric acetylcholine receptors when expressed in host cells.

36. (New) A process for preparing a polypeptide encoded by a nucleic acid of Claim 35 comprising

(a) culturing a host cell comprising at least one nucleic acid of Claim 35, wherein the nucleic acid is functionally linked to regulatory sequences which ensure expression of the nucleic acid in prokaryotic or eukaryotic cells, and

(b) isolating the polypeptide from the cell or culture medium.

ATTACHMENT 1

Appendix 1
Pairwise alignments of SEQ ID NOs 1,3,5

The pairwise alignments were performed using the Bestfit program from the Wisconsin Package with standard parameters. The following sequence ranges were used for the Bestfit algorithm: SEQ ID NO1: from position 1295 to position 2195, SEQ ID NO3 from position 432 to position 1318, and SEQ ID NO5 from position 154 to position 1123. Bestfit will only display the regions of greatest identity.

BESTFIT of: seqid5.seq check: 5528 from: 1 to: 980

SEQID5

to: seqid1.seq check: 9365 from: 1 to: 901

SEQID1

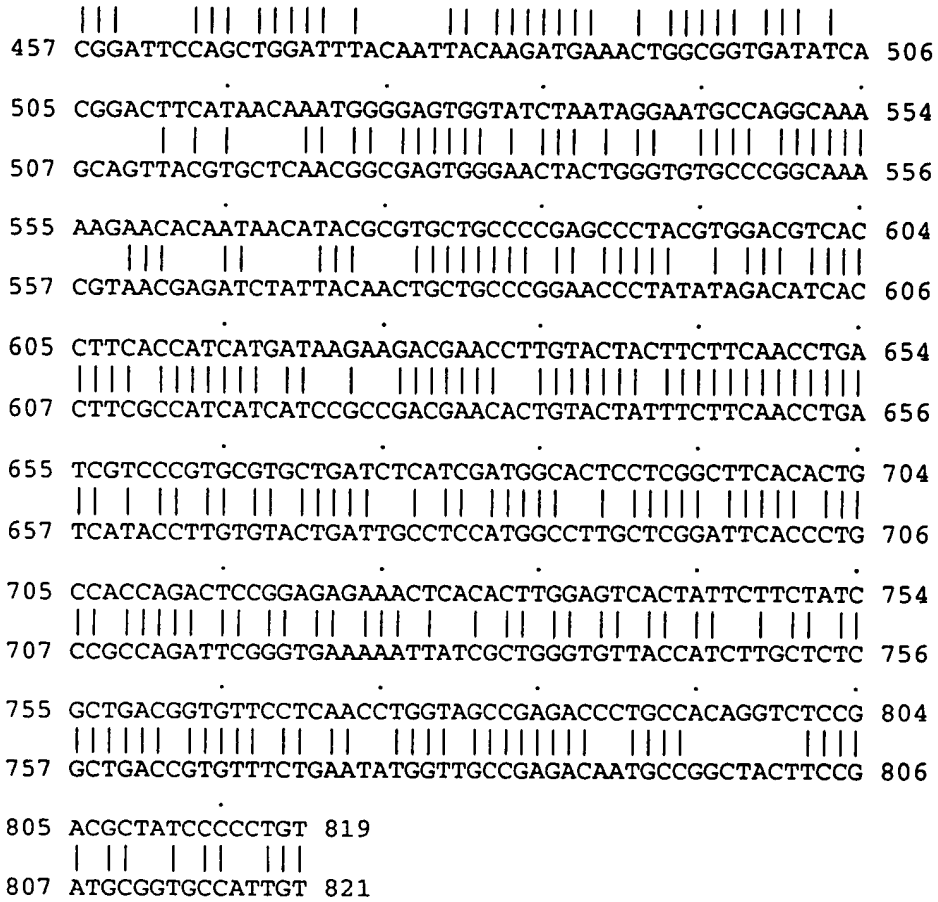
Symbol comparison table: /r4/sftw/GCGHOME/gcgcore/data/rundata/swgapdna.cmp
CompCheck: 2335

Gap Weight:	50	Average Match:	10.000
Length Weight:	3	Average Mismatch:	-9.000
Quality:	3343	Length:	815
Ratio:	4.102	Gaps:	0
Percent Similarity:	68.957	Percent Identity:	68.957

Match display thresholds for the alignment(s):
| = IDENTITY
: = 5
. = 1

seqid5.seq x seqid1.seq October 9, 2001 17:29 ..

```
          .         .         .         .         .
  5 TCACGAGAAGAGACTCCTGAACGCGTTGCTGGCGAACTACAACACCCTGG 54
    ||| || ||||||||  |  |||  |  |||    || || || || |
  7 TCATGAAAAGAGACTGTTACACGATCTTTTGGATCCTTATAATACACTAG 56
          .         .         .         .         .
 55 AGCGACCGGTGGCCAACGAGAGCGAACCGCTAGAGGTCAGGTTCGGCTTG 104
    | || || ||   ||| ||    || ||| || |  | || || || ||
 57 AACGTCCCGTTCTCAATGAATCGGACCCGTTACAATTAAGCTTTGGTTTA 106
          .         .         .         .         .
105 ACCTTGCAGCAAATCATTGACGTGGACGAGAAGAATCAACTACTTATAAC 154
    || ||   |||||| || || |||||||||||| |||||| | ||  | ||
107 ACTTTAATGCAAATTATCGATGTGGACGAGAAAAATCAATTGCTAGTCAC 156
          .         .         .         .         .
155 CAATATATGGCTGTCGTTGGAGTGGAATGACTACAACCTGAGGTGGAACG 204
     ||| | ||| |     ||||||||||| |||   || ||  | ||||||
157 TAATGTGTGGTTAAAACTGGAGTGGAACGACATGAATCTCCGCTGGAACA 206
          .         .         .         .         .
205 ACAGCGAGTATGGCGGGGTCAAGGACCTCAGGATCACGCCCAACAAGTTG 254
     |  ||| |||||||| || ||||| ||  | ||  |||| |     |
207 CCTCCGACTATGGCGGAGTTAAGGATCTGCGAATACCGCCGCATCGCATC 256
          .         .         .         .         .
255 TGGAAGCCGGACGTCCTTATGTATAATAGTGCTGACGAGGGTTTTGACGG 304
    |||||||||||||| || ||||| || ||||| || ||||| ||||||||
257 TGGAAGCCGGACGTGCTGATGTACAACAGTGCGGATGAGGGATTTGACGG 306
          .         .         .         .         .
305 GACCTACCAGACCAACGTGGTGGTCAGAAGCGGCGGCAGTTGCCTGTACG 354
     |||||||||||| ||||||||||  | | |  ||||   || || ||||
307 CACCTACCAGACGAACGTGGTGGTGCGGAACAACGGCTCGTGTCTATACG 356
          .         .         .         .         .
355 TGCCACCTGGCATATTCAAGAGCACATGCAAGATGGACATCGCGTGGTTT 404
    | || || || || ||||||   || |||||||| |||||| |||||||
357 TTCCGCCGGGGATCTTCAAGTCGACGTGCAAGATCGACATCACGTGGTTC 406
          .         .         .         .         .
405 CCCTTCGACGACCAACACTGTGATATGAAGTTCGGTAGCTGGACATATGA 454
    |||||||| ||||| |  || || |||||||||| || ||||| || ||
407 CCCTTCGATGACCAGCGGTGCGAGATGAAGTTCGGCAGTTGGACCTACGA 456
          .         .         .         .         .
455 CGGCAATCAGTTGGATCTGGTGCTAAAAGATGAGGCAGGCGGCGATCTAT 504
```



BESTFIT of: seqid1.seq check: 9365 from: 1 to: 901

SEQID1

to: seqid3.seq check: 9434 from: 1 to: 887

SEQID3

Symbol comparison table: /r4/sftw/GCGHOME/gcgcore/data/rundata/swgapdna.cmp
CompCheck: 2335

Gap Weight:	50	Average Match:	10.000
Length Weight:	3	Average Mismatch:	-9.000
Quality:	4111	Length:	812
Ratio:	5.063	Gaps:	0
Percent Similarity:	74.015	Percent Identity:	74.015

Match display thresholds for the alignment(s):
| = IDENTITY
: = 5
. = 1

seqid1.seq x seqid3.seq October 23, 2001 17:36 ..

```
          .         .         .         .         .
  9 ATGAAAAGAGACTGTTACACGATCTTTTGGATCCTTATAATACACTAGAA 58
    | || ||| | ||  | ||| | || ||||| |  || ||   ||| ||
  1 ACGAGAAGCGGCTACTGCACCACCTATTGGACCACTACAACGTACTGGAG 50
          .         .         .         .         .
 59 CGTCCCGTTCTCAATGAATCGGACCCGTTACAATTAAGCTTTGGTTTAAC 108
     | |||||  |||| ||    |||||| | ||  |   ||| ||  | ||
 51 AGGCCCGTCGTCAACGAGAGCGACCCGCTGCAGCTCTCCTTCGGCCTCAC 100
          .         .         .         .         .
109 TTTAATGCAAATTATCGATGTGGACGAGAAAAATCAATTGCTAGTCACTA 158
     | ||||| || ||||| |||||||||||| || ||  |  || | || |
101 GCTCATGCAGATCATCGACGTGGACGAGAAGAACCAGCTTTTAATAACAA 150
          .         .         .         .         .
159 ATGTGTGGTTAAAACTGGAGTGGAACGACATGAATCTCCGCTGGAACACC 208
    |  | ||| ||||||| |||||||| || |||||  |  | ||||||||
151 ACATCTGGCTAAAACTAGAGTGGAATGATATGAACTTGAGGTGGAACACT 200
          .         .         .         .         .
209 TCCGACTATGGCGGAGTTAAGGATCTGCGAATACCGCCGCATCGCATCTG 258
    || || |  ||||| || || ||| |  || | || || ||  |  | ||
201 TCAGATTTCGGCGGGGTCAAAGATTTAAGAGTGCCACCCCACAGACTATG 250
          .         .         .         .         .
259 GAAGCCGGACGTGCTGATGTACAACAGTGCGGATGAGGGATTTGACGGCA 308
    ||| || ||||| || |||||||||| ||||| || || || ||| |||
251 GAAACCAGACGTCCTTATGTACAACAGCGCGGACGAAGGGTTCGACAGCA 300
          .         .         .         .         .
309 CCTACCAGACGAACGTGGTGGTGCGGAACAACGGCTCGTGTCTATACGTT 358
    | || |  ||||||||||||||||||||||||||||||||||| |||||
301 CGTATCCAACGAACGTGGTGGTGCGGAACAACGGCTCGTGTCTGTACGTG 350
          .         .         .         .         .
359 CCGCCGGGGATCTTCAAGTCGACGTGCAAGATCGACATCACGTGGTTCCC 408
    ||||| || |||||||||   || ||||||||||||||||| ||||||||
351 CCGCCCGGCATCTTCAAGAGCACCTGCAAGATCGACATCACCTGGTTCCC 400
          .         .         .         .         .
409 CTTCGATGACCAGCGGTGCGAGATGAAGTTCGGCAGTTGGACCTACGACG 458
    |||||| ||||| || |||||||||||||| ||||| ||||| || || |
401 CTTCGACGACCAACGATGCGAGATGAAGTTTGGCAGCTGGACTTATGATG 450
          .         .         .         .         .
459 GATTCCAGCTGGATTTACAATTACAAGATGAAACTGGCGGTGATATCAGC 508
```

```
451 GTTATCAGTTGGATCTACAACTACAGGATGAAGGGGGCGGAGATATAAGC 500

509 AGTTACGTGCTCAACGGCGAGTGGGAACTACTGGGTGTGCCCGGCAAACG 558
501 AGTTTTGTCACGAATGGCGAATGGGAGTTAATAGGAGTCCCCGGCAAGCG 550

559 TAACGAGATCTATTACAACTGCTGCCCGGAACCCTATATAGACATCACCT 608
551 CAACGAGATCTACTACAACTGTTGTCCGGAGCCATACATCGACATCACGT 600

609 TCGCCATCATCATCCGCCGACGAACACTGTACTATTTCTTCAACCTGATC 658
601 TTGCGGTGGTGATCCGGAGGAAAACGCTCTACTACTTCTTCAATCTGATC 650

659 ATACCTTGTGTACTGATTGCCTCCATGGCCTTGCTCGGATTCACCCTGCC 708
651 GTGCCCTGCGTGCTCATCGCCTCCATGGCTCTATTGGGGTTCACCTTGCC 700

709 GCCAGATTCGGGTGAAAAATTATCGCTGGGTGTTACCATCTTGCTCTCGC 758
701 TCCAGACTCCGGAGAAAAGTTGTCTTTAGGTGTGACGATATTACTGTCGT 750

759 TGACCGTGTTTCTGAATATGGTTGCCGAGACAATGCCGGCTACTTCCGAT 808
751 TGACGGTGTTCCTCAACATGGTGGCGGAGACGATGCCAGCGACGTCGGAC 800

809 GCGGTGCCATTG 820
801 GCCGTGCCCTTG 812
```

BESTFIT of: seqid5.seq check: 5528 from: 1 to: 980

SEQID5

to: seqid3.seq check: 9434 from: 1 to: 887

SEQID3

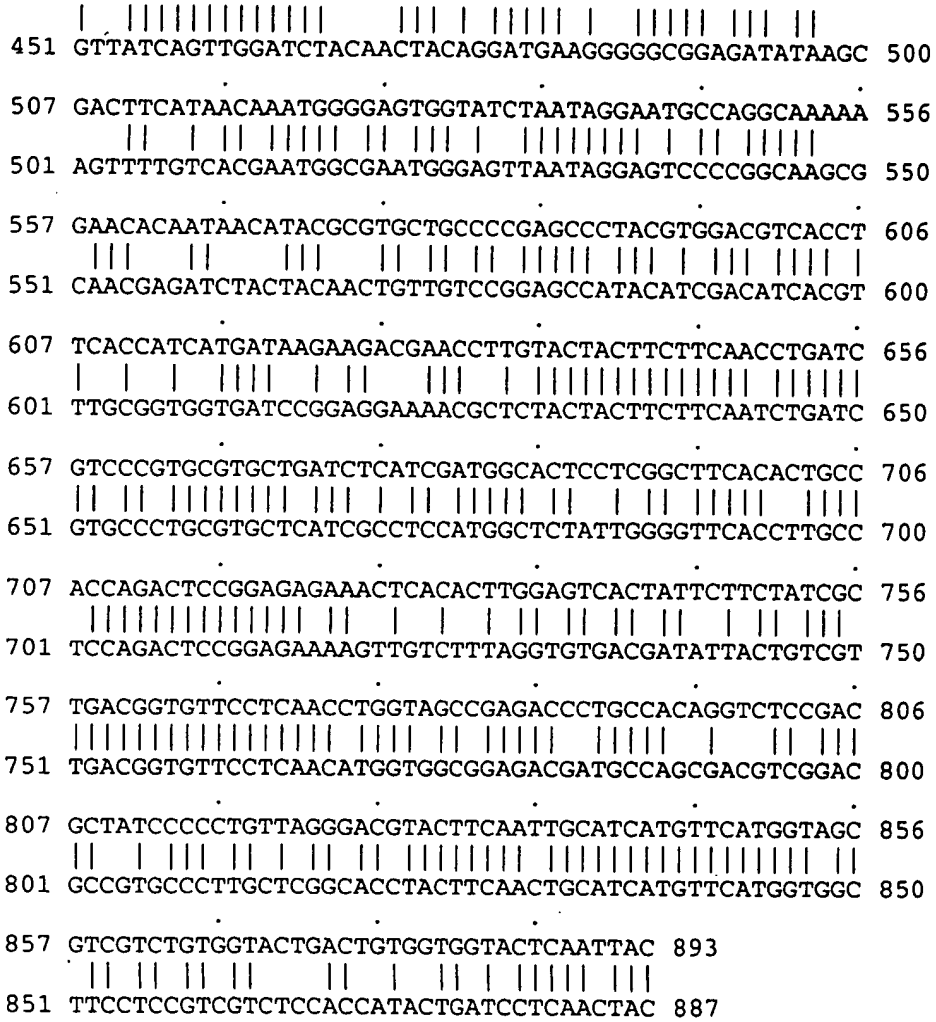
Symbol comparison table: /r4/sftw/GCGHOME/gcgcore/data/rundata/swgapdna.cmp
CompCheck: 2335

Gap Weight:	50	Average Match:	10.000
Length Weight:	3	Average Mismatch:	-9.000
Quality:	4006	Length:	887
Ratio:	4.516	Gaps:	0
Percent Similarity:	71.139	Percent Identity:	71.139

Match display thresholds for the alignment(s):
| = IDENTITY
: = 5
. = 1

seqid5.seq x seqid3.seq October 9, 2001 17:31 ..

```
      7 ACGAGAAGAGACTCCTGAACGCGTTGCTGGCGAACTACAACACCCTGGAG 56
        |||||||| | || ||| ||    |  |||   ||||||||   ||||||
      1 ACGAGAAGCGGCTACTGCACCACCTATTGGACCACTACAACGTACTGGAG 50

     57 CGACCGGTGGCCAACGAGAGCGAACCGCTAGAGGTCAGGTTCGGCTTGAC 106
         | || || | ||||||||||||| |||||  || ||   ||||||| | ||
     51 AGGCCCGTCGTCAACGAGAGCGACCCGCTGCAGCTCTCCTTCGGCCTCAC 100

    107 CTTGCAGCAAATCATTGACGTGGACGAGAAGAATCAACTACTTATAACCA 156
          |   ||| ||||| |||||||||||||||||| || ||  | ||||| |
    101 GCTCATGCAGATCATCGACGTGGACGAGAAGAACCAGCTTTTAATAACAA 150

    157 ATATATGGCTGTCGTTGGAGTGGAATGACTACAACCTGAGGTGGAACGAC 206
        | || |||||     | ||||||||||    ||| ||||||||||
    151 ACATCTGGCTAAAACTAGAGTGGAATGATATGAACTTGAGGTGGAACACT 200

    207 AGCGAGTATGGCGGGGTCAAGGACCTCAGGATCACGCCCAACAAGTTGTG 256
           || |  |||||||||| ||  | ||  |  | ||| |||   | ||
    201 TCAGATTTCGGCGGGGTCAAAGATTTAAGAGTGCCACCCCACAGACTATG 250

    257 GAAGCCGGACGTCCTTATGTATAATAGTGCTGACGAGGGTTTTGACGGGA 306
        ||| || |||||||||||||| || || || ||||| || || ||| | |
    251 GAAACCAGACGTCCTTATGTACAACAGCGCGGACGAAGGGTTCGACAGCA 300

    307 CCTACCAGACCAACGTGGTGGTCAGAAGCGGCGGCAGTTGCCTGTACGTG 356
        | || |  || ||||||||||  | | |  ||||   || |||||||||
    301 CGTATCCAACGAACGTGGTGGTGCGGAACAACGGCTCGTGTCTGTACGTG 350

    357 CCACCTGGCATATTCAAGAGCACATGCAAGATGGACATCGCGTGGTTTCC 406
        || || ||||| |||||||||| |||||||| |||||| | ||||| ||
    351 CCGCCCGGCATCTTCAAGAGCACCTGCAAGATCGACATCACCTGGTTCCC 400

    407 CTTCGACGACCAACACTGTGATATGAAGTTCGGTAGCTGGACATATGACG 456
        |||||||||||||  || || |||||||| || |||||||| ||||| |
    401 CTTCGACGACCAACGATGCGAGATGAAGTTTGGCAGCTGGACTTATGATG 450

    457 GCAATCAGTTGGATCTGGTGCTAAAAGATGAGGCAGGCGGCGATCTATCG 506
```



ATTACHMENT 2

Appendix 2
Multiple alignment of nicotinic acetylcholine receptor alpha subunits

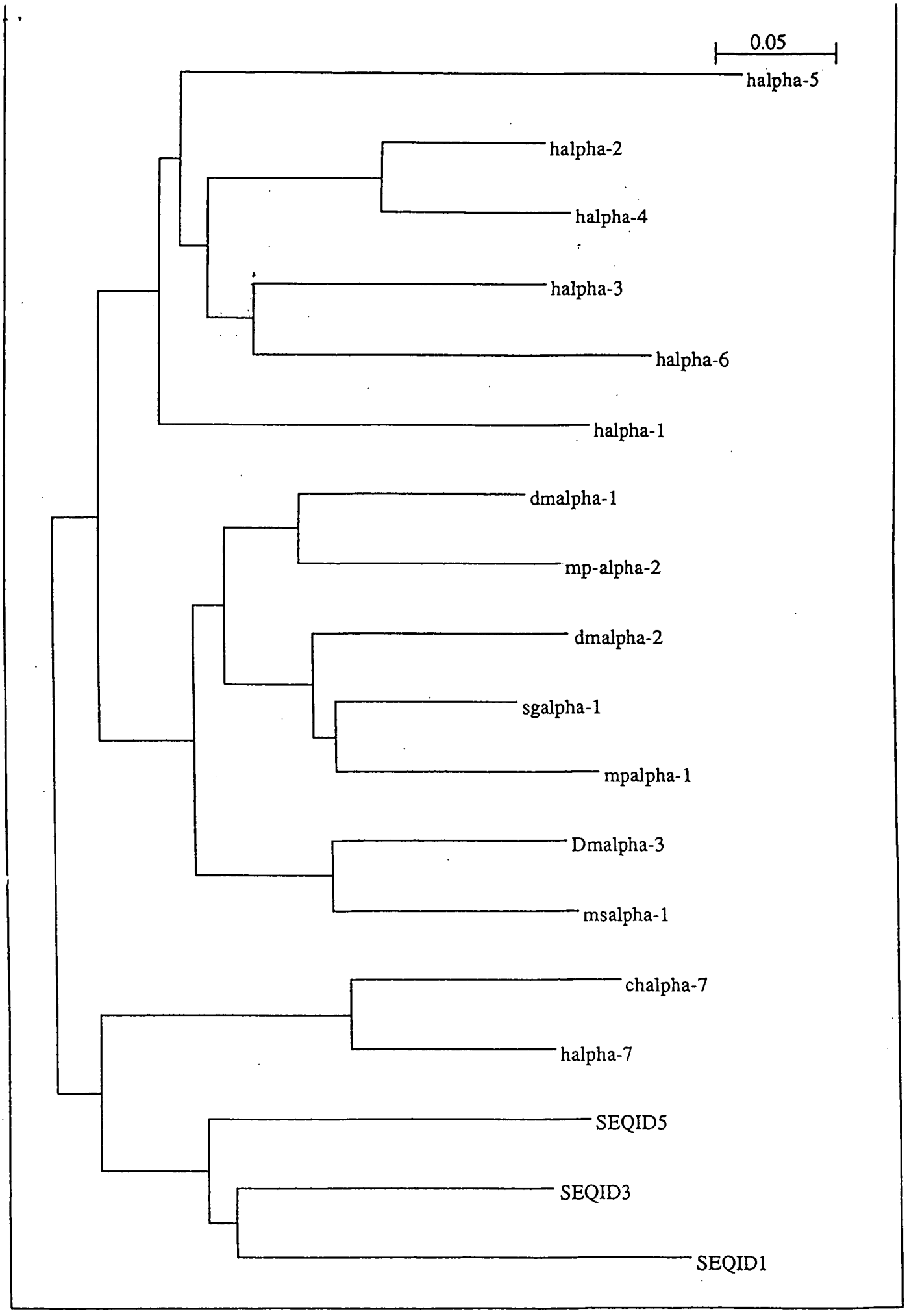
The following sequences were multiply aligned using the program ClustalX 1.8 with standard settings:

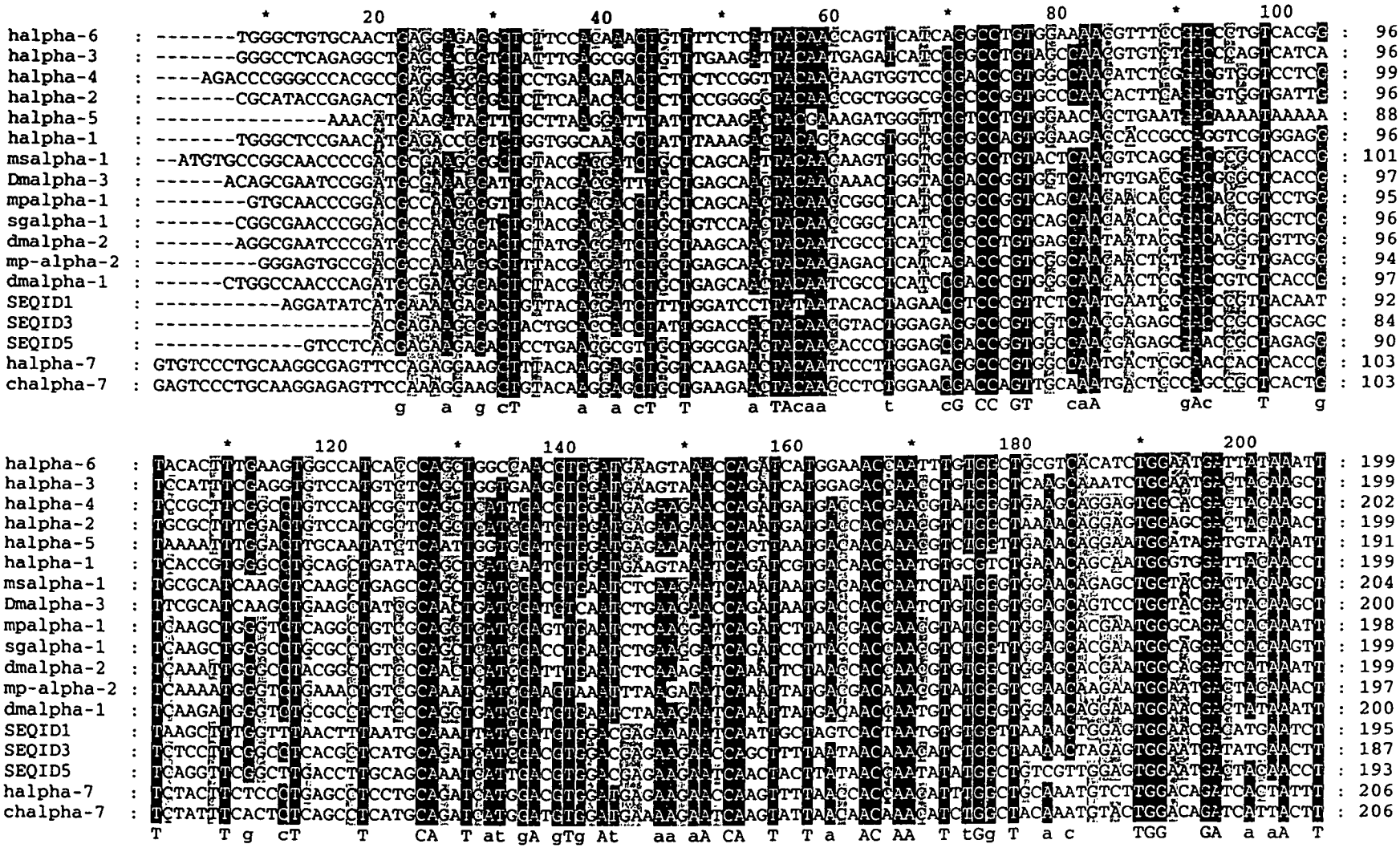
Mouse α-7: gi6671502
Chick α-7: gi287756 (chalpha-7)
Human α-7: gi4502830 (halpha-7)
Manduca sexta alpha-1: Y09795 (msalpha-1)
Myzus persicae alpha-1: X81887 (mpalpha-1)
Myzus persicae alpha-2: X81888 (mpalpha-2)
Schistocerca gregaria alpha-1: X55439 (sgalpha-1)
Drosophila alpha-1: X07194 (dmalpha-1)
Drosophila alpha-2: X52274 (dmalpha-2)
Drosophila alpha-3: Y15593 (dmalpha-3)
Human alpha-1: XM_027441 (halpha-1)
Human alpha-2: XM_027441 (halpha-2)
Human alpha-3: XM_039145 (halpha-3)
Human alpha-4: XM_028857 (halpha-4)
Human alpha-5: XM_039143 (halpha-5)
Human alpha-6: XM_005283 (halpha-6)

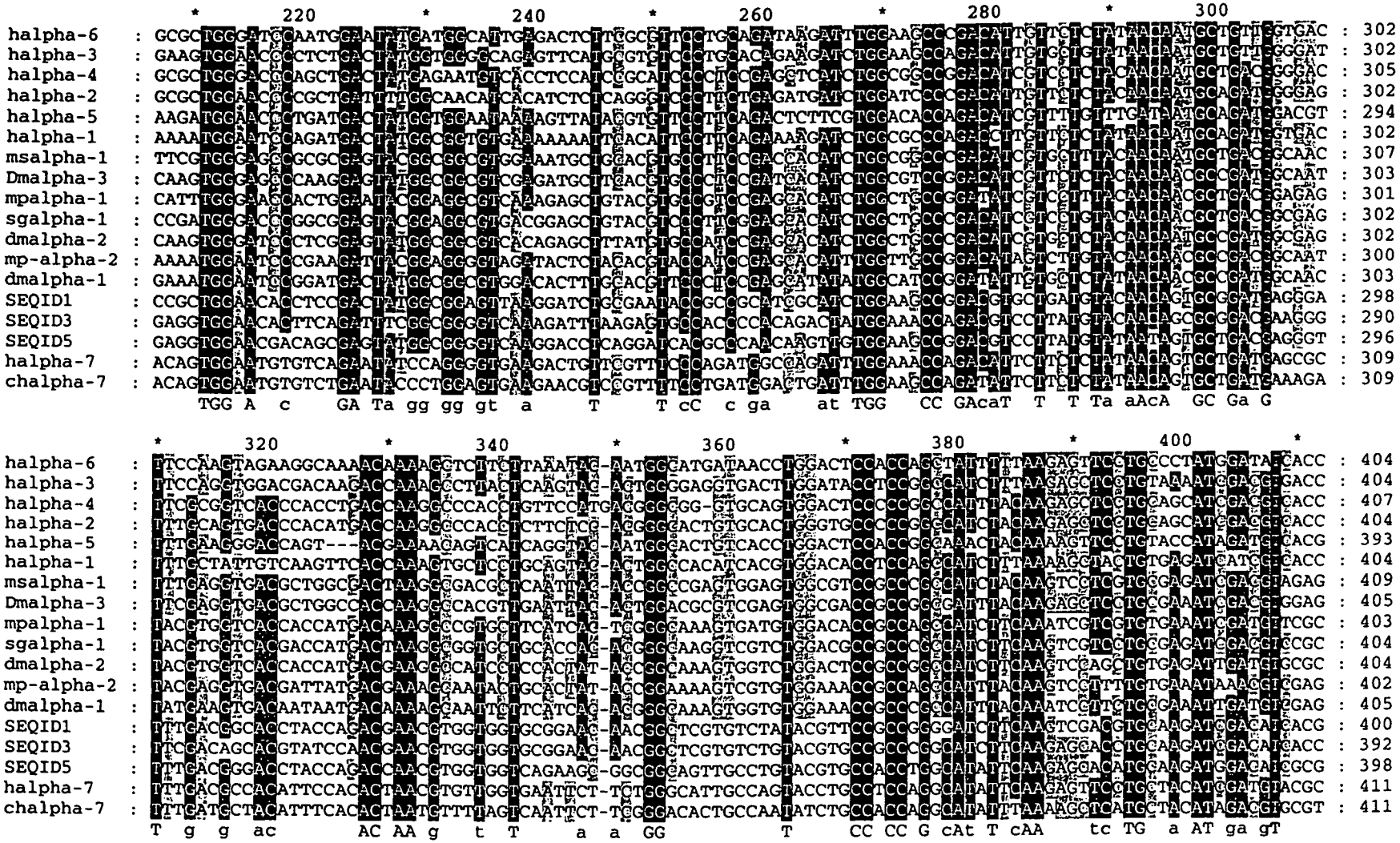
Sequence ranges of SEQIDs NO1,3,5:
SEQ ID NO1: from position 1295 to position 2195,
SEQ ID NO3 from position 432 to position 1318, and
SEQ ID NO5 from position 154 to position 1123.

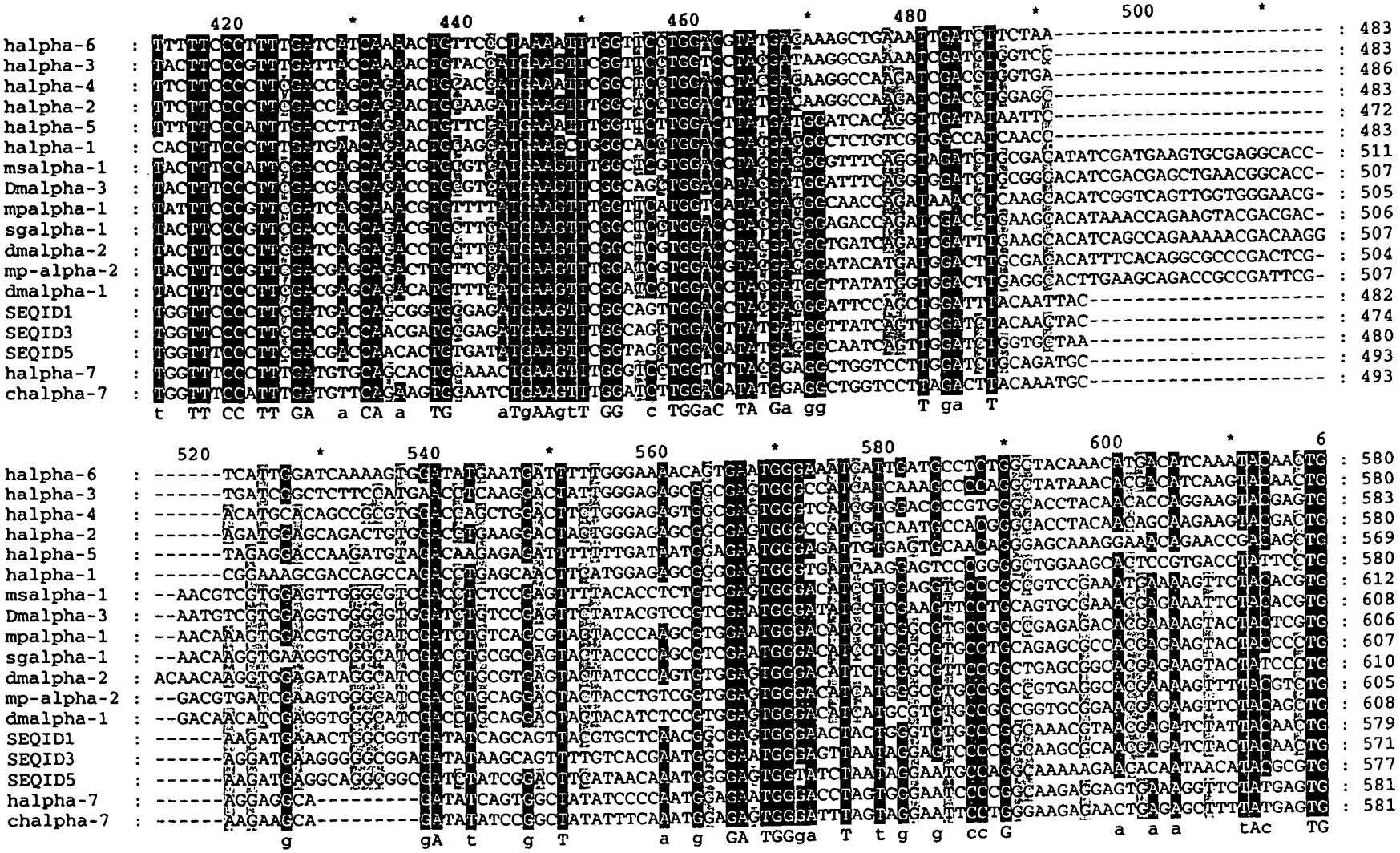
The ClustalX program is described in: Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997); The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research, 24:4876-4882.)

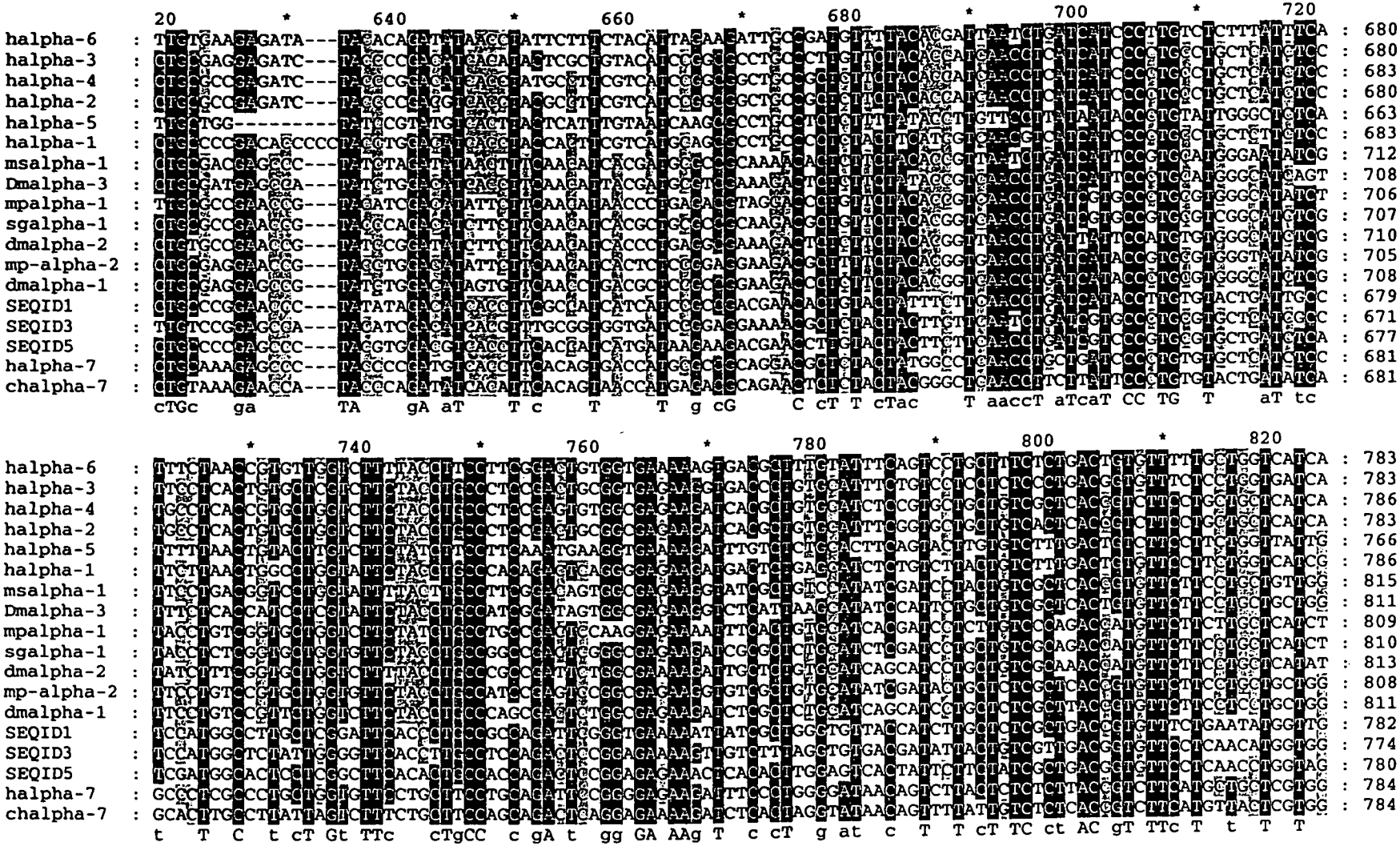
The output of the program consists of the multiple alignment (included) and a dendrogram (diagram indicating relatedness).The dendrogram output file from the multiple alignment program ClustalX was processed by the program Njplot which is part of the ClustalX program package.

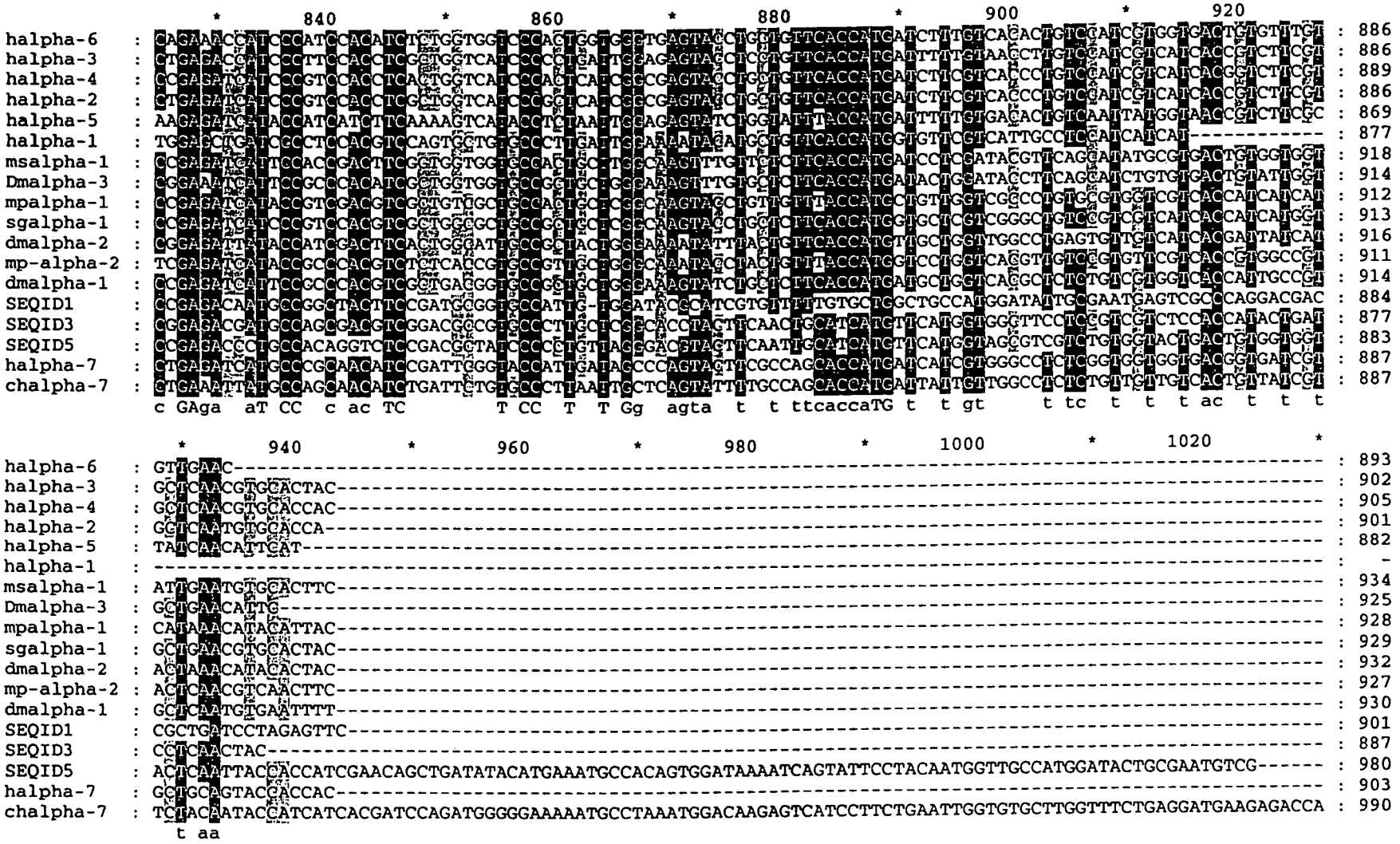

0.05
halpha-5
halpha-2
halpha-4
halpha-3
halpha-6
halpha-1
dmalpha-1
mp-alpha-2
dmalpha-2
sgalpha-1
mpalpha-1
Dmalpha-3
msalpha-1
chalpha-7
halpha-7
SEQID5
SEQID3
SEQID1




220
240
260
280
300
halpha-6 : GCGCTGGGATGCAATGGAATTTCATGGCATTGAGACTCTTGGCGTTCCTGCAGATAAGATTTGGAAGCCCGACATTGTTCTCTTTAACAATGCTGTTGGTGAC : 302
halpha-3 : GAAGTGGAACCCCTCTGACTATGGTGGGCAGAGTTCATGCGTGTCCCTGCACAGAAGATCTGGAAGCCAGACATTGTGCTGTTTAACAATGCTGTTGGGGAT : 302
halpha-4 : GCGCTGGGACGCAGCTGACTTTCAGAATGTCACCTCCATCGGCATCCCCTCCGAGGTCATCTGGCGCCGGACATCGTCCTCTACAACAATGCTGACGGGAC : 305
halpha-2 : GCGCTGGAACGCCGCTGATTTTGGCAACATCACATCTCTCAGGGTCCCTTCTGAGATGATCTGGATCCCGACATTGTTGTCTACAACAATGCAGATGGGAG : 302
halpha-5 : AAGATGGAACCCTGATGACTATGGTGCAATAATAGTTATAGGTGTTCCTTCAGACTCTTCGTGGACACCAGACATCGTTTTGTTTGTTAATGCAGATGGACGT : 294
halpha-1 : AAAATGGAATCCAGATGACTATGGCGGTGTGAAAAAATTCACATTCCTTCAGAAAAGATCTGGCGCCCAGACCTTGTTGTCTATAACAATGCAGATGGTGAC : 302
msalpha-1 : TTCGTGGGAGCCGCGCGAGTACGGCGGCGTGGAAATGCTGCACGTGCCTTCCGACCACATCTGGCGCCCGACATCGTGGTTTACAACAATGCTGACGCAAC : 307
Dmalpha-3 : CAAGTGGGAGCCCAAGGAGTATGGCGGCGTCGAGATGCTTCACGTGCCCTCCGATCACATCTGGCGTCCGGACATCGTTCTCTTCAACAACGCCGATGGCAAT : 303
mpalpha-1 : CATTTGGGAAGCACTGGAATACGGCAGGCGTCAAAGAGCTGTACGTGCCGTCCGAGCACATCTGGCTGCCGGATATCGTCGTTTACAACAACGCTGACGGAGAG : 301
sgalpha-1 : CCGATGGGACCCGGCGGAGTACGGCAGGCGTGACGGAGCTGTACGTCCCTTCGGAGCACATCTGGCTGCCCGACATCGTCGTGTACAACAATGCTGACGGCGAG : 302
dmalpha-2 : CAAGTGGGATCCCTCGGAGTATGGCGGCGTCACAGAGCTTTATGTGCCATCCGAGCACATCTGGCTGCCCGACATCGTGGTCTACAACAATGCCGATGGCAAT : 302
mp-alpha-2 : AAAATGGAATCCCGAAGATTACGCAGGGGTAGATACTCTACACGTACCATCCGAGCACATTTGGTTGCCGGACATAGTCTTGTTACAACAACGCCGACGGCAAT : 300
dmalpha-1 : GAAATGGAATCCGGATGACTATGGCGGCGTGGACACTTTGCACGTTCCCTCCGAGCATATATGGCATCCGGATATTTGTGCTCTTTAACAACGCCGATGGCAAC : 303
SEQID1 : CCGCTGGAACACCTCCGACTATGGCCGCAGTTAAGGATCTGCACGTTACCGCCGCATCGCATCTGGAAGCCGGACGTGCTGATGTACAACAGTGCGGCGATCAGGGA : 298
SEQID3 : GAGGTGGAACACTTCAGATTTCGGCGCGGTCAAAGATTTAAGAGTGCCACCCCACAGACTATGGAAACCAGACGTCCTTATGTACAACAGCGCGGACCAAGGG : 290
SEQID5 : GAGGTGGAACGACAGCGAGTATGGCGCGGGTCAAGGACCTCAGGATCACGCCCAACAAGTGTGTGGAAGCCGGACGTCCTTATGTGTAATAGTGCTGACCAGGGT : 296
halpha-7 : ACAGTGGAATGTGTCAGAATATCCAGGGGTGAAGACTGTTCGTTTCCCAGATGGCCAGATTTGGAAACCAGACATTCTTCTCTATAACAGTGCTGATCAGCGC : 309
chalpha-7 : ACAGTGGAATGTGTCTGAATACCCTGCAGTGAAGAACGTCCGTTTTCCTGATGGACTGATTTGGAAGCCAGATATTCTTCTCTATAACAGTGCTGATCAAAGA : 309
TGG A c GA Ta gg gg gt a T T cC c ga at TGG CC GAcaT T T Ta aAcA GC Ga G
320
340
360
380
400
halpha-6 : TTCCAAGTAGAAGGCAAAACAAAAGCTCTTCTTAAATAC-AATGGCATGATAACCTGGACTCCACCAGCTATTTTTAAGAGTTCGTGCCCTATGGATAACACC : 404
halpha-3 : TTCCAGGTGGACGACAAGACCAAAGCCTTACTCAAGTAC-ACTGCGGAGGTGACTTGGATACCTCCGGCCATCTTTAAGAGCTCGTGTGTAAAATGGAGCGGACC : 404
halpha-4 : TTCCCGGTCACCCACCTGACCAAGGCCCCACCTGTTCCATGAGGGCGG-GTGCAGTGGACTCCCCCGGCCATTTACAAGAGCTCGTGCAGCATCGAGCGTCACC : 407
halpha-2 : TTTGCAGTGACCCACATGACCAAGGCCCACCTCTTCTCC-ACGGGCACTGTGCACTGGGTGCCCCCGGCCATCTACAAGAGCTCGTGCAGCATCGAGCGTCACC : 404
halpha-5 : TTTGAAGGGACCAGT---ACGAAAACAGTCATCAGGTAC-AATGGCACTGTCACCTGGACTCCACCGGCAAACTACAAAAGTTCCTGTACCATAGATGTCACG : 393
halpha-1 : TTTGCTATTGTCAAGTTCACCAAAGTGCTCCTGCAGTAC-ACTGGCCACATCACGTGGACACCTCCAGCCATCTTTAAAAGCTACTGTGAGATCATCGCCACC : 404
msalpha-1 : TTTGAGGTGACGCTGGCGACTAAGGCGACGGTCAATTAC-ACCGGCCGAGTGGAGTGGCGTCCGCCCGCCATCTACAAGTCCTCGTGCGAGATCGACGTAGAG : 409
Dmalpha-3 : TTCGAGGTGACGCTGGCCACCAAGGCCACGTTGAATTAC-ACTGCACGCGTCGAGTGGCGACCGCCGGCGATTTACAAGAGCTCGTGCGAAATCGACGTGGAG : 405
mpalpha-1 : TACGTGGTCACCACCATGACAAAGGCCGTGCTTCATCAC-TGCGGCAAAGTGATGTGGACACCGCCAGCCATCTTCAAATCGTCGTGTGAAATCGATGTTCGC : 403
sgalpha-1 : TACGTGGTCACGACCATGACTAAGGCGGTGCTGCACCAC-ACGGGCAAGGTCGTCTGGACGCCGCCCGCCATCTTCAAGTCGTCGTGCGAGATCGACGTCGCC : 404
dmalpha-2 : TACGTGGTCACCACCATGACGAAGGCCATCCTCCACTAT-ACCGCCAAAGTGGTCTGGACTCCGCCGGCCATCTTCAAGTCCAGCTGTGAGATTGATGTGCGC : 404
mp-alpha-2 : TACGAGGTGACGATTATGACGAAAGCAATACTGCACTAT-ACCGCGAAAAGTCGTGTGTGGAAACCGCCCAGGCATTTACAAGTCCTTTGTGAAATAAAGGTCGAG : 402
dmalpha-1 : TATGAAGTGACAATAATGACAAAAGCAATTGTTCATCAG-ACGGGCAAAGTGGTGTGGAAACCGCCCGCCATTTACAAATCGTTGTGCGAAATTGATGTCGAG : 405
SEQID1 : TTTCACGGCACCTACCAGACGAACGTGGTGGTGCCGGAAC-AACGGCTCGTCTGTCTATACGTTCCGCCGGGGATCTTCAAGTCGACGTGCAAGATGGACATGACG : 400
SEQID3 : TTCGACAGCACGTATCCAACGAACGTGGTGGTGCGGAAC-AACGGCTCGTGTCTGTACGTGCCGCCCGGCATCTTCAAGAGGACCTGCAAGATGGACATGACC : 392
SEQID5 : TTTGACGGGACCTACCAGACCAACGTGGTGGTCAGAAGC-GGCGGCAGTTGCCTGTACGTGCCACCTGGCATATTCAAGAGGACATGCAAGATGGAGATCGCG : 398
halpha-7 : TTTGACGCCACATTCCACACTAACGTGTTGGTGAATTCT-TCTGGCATTGCCAGTACCTGCCTCCAGGCATATTCAAGAGTTCGTGCTACATGGATGTACGC : 411
chalpha-7 : TTTCATGCTACATTTCACACTAATGTTTTAGTCAATTCT-TGCGGCACACTGCCAATATCTGCCACCAGGCATATTTAAAAGCTCATGCTACATAGACGTGCGT : 411
T g g ac AC AA g t T a a GG T CC CC G cAt T cAA tc TG a AT ga gT


420 * 440 * 460 * 480 * 500 *
halpha-6 : 483
halpha-3 : 483
halpha-4 : 486
halpha-2 : 483
halpha-5 : 472
halpha-1 : 483
msalpha-1 : 511
Dmalpha-3 : 507
mpalpha-1 : 505
sgalpha-1 : 506
dmalpha-2 : 507
mp-alpha-2 : 504
dmalpha-1 : 507
SEQID1 : 482
SEQID3 : 474
SEQID5 : 480
halpha-7 : 493
chalpha-7 : 493
520 * 540 * 560 * 580 * 600 * 6
halpha-6 : 580
halpha-3 : 580
halpha-4 : 583
halpha-2 : 580
halpha-5 : 569
halpha-1 : 580
msalpha-1 : 612
Dmalpha-3 : 608
mpalpha-1 : 606
sgalpha-1 : 607
dmalpha-2 : 610
mp-alpha-2 : 605
dmalpha-1 : 608
SEQID1 : 579
SEQID3 : 571
SEQID5 : 577
halpha-7 : 581
chalpha-7 : 581


20 * 640 * 660 * 680 * 700 * 720
halpha-6 : TTGTGAAGAGATA---TACACACGTTAACCTATTCTTTCTACATTACAAGATTGCCGATGTTTTACACGATTAATTGTGATGTTCCCTTGTCTCTTTATTTCA : 680
halpha-3 : CTGCGAGGAGATC---TACCCCGACATCACATAGTCGCTGTACATCCGGCGCCTGCCCTTGTTCTACACCATGAACGTCATCATCCCGTGCCTGCTCATGTCC : 680
halpha-4 : CTGCGCCGAGATC---TACCCGGACATCACGATGCCTTCGTCATCCGGCGGCTTCCGCGTGTTCTACACCATGAACGCTCATCATCCCGTGCCTGCTCATGTCC : 683
halpha-2 : CTGCGCCGAGATC---TACGCCCGAGGTCACGTACGCGTTCGTCATCCGGCGGCTGCCGCGTGTTCTACACCATGAACGCTCATCATCCCGTGCCTGCTCATGTCC : 680
halpha-5 : TTGCTGG---------TATTCGTATGTCACTTACTCATTTGTAATCAAGCGCCTGCCTCTGTTTTATACCTTGTTCGTTATAATACCGTGTAATTGGGCTGTCA : 663
halpha-1 : CTGCCCCGACACGCCCCTACGTGCAGTTCACGTACCACTTCGTCATGCAGCGCCTGCCCGTGTACTTCATGGTCAACGTCATCATCCCGTGCCTGCTCTTGTCC : 683
msalpha-1 : CTGCGACGAGCCC---TATCTACGTTTAACTTTCAACATCACGATGGCCGCCAAAACACGTCTTCTACACCGTTAATCTGATCATTCCGTGCATGGGAATATCG : 712
Dmalpha-3 : CTGCGATGAGCCA---TATCTGCACATCACGTTCAACATTACGATGCGTCGAAAGACTCTGTTCTATACCGTGAACCTGCTCATTCCGTGCATGGGCATCAGT : 708
mpalpha-1 : TTGCGCCGAACCG---TACATCCACATATTCTTCAAGATAACCCTGAGACGTAGCACCGTGTTCTACACGGTGAACCTGCTCGTGCCGTGCGTGGGCATATCT : 706
sgalpha-1 : CTGCGCCGAAGCG---TACCCACACATCTTCTTCAACATCACGCTGCCGCCGCAAGACGCGTTCTACACGGTGAACCTGCTTATTCCATGTGTGGGCATCTCG : 707
dmalpha-2 : CTGTGCCGAACCG---TATCCCGCATATCTTCTTCAACATCACCCTGAGGGGAAACACTCTGTTCTACACGGTGAACCTGCTTATTCCATGTGTGGGCATCTCG : 710
mp-alpha-2 : CTGCGAGGAACCG---TAGGTGCACATATTCTTCAACATCACTCTCCGGAGGAAGACGCGTTTTCTACACGGTGAACCTGCTCATTCCGTGCGTGGGCATCTCG : 705
dmalpha-1 : CTGCGAGGAGCCG---TATCTGCACATAGTGTTCAACCTGACGCTCCGCCGGAAGACCCTGTTCTACACGGTGAACCTGCTCATTCCGTGCGTGGGCATCTCG : 708
SEQID1 : CTGCCCCGGAACCC---TATATACACATCACGTTGCGCATCATCATCCGCCGACGAACACGTGTACTATTTCTTGAACCTGCTCATTCCTTGTGTACTGATTGCC : 679
SEQID3 : TTGTCCGGAGCCA---TACATCCACATCACGTTTGCGGTGGTGATCCGGAGGAAAACGCGTCTACTACTTGTTCAACTGCTCATCGTGCCGTGCGTGCTCATCGCC : 671
SEQID5 : CTGCCCCGAGCCC---TACGTGCACGTCACGTTCACGATCATGATAAGAAGACGAACCTTGTACTACTTCTTCAACCTGCTCATCGTCCCGTGCGTGCTGATCTCA : 677
halpha-7 : CTGCAAAGAGCCC---TACCCCCACTGTCACCTTCACAGTGACCATGGCCGCAGGACGCGTCTACTATGGCCTGAACCTGCTGATCCCCTGTGTGTGCTCATCTCC : 681
chalpha-7 : CTGTAAAGAACCA---TACCCACACTATCACAATTCACAGTAACCATGACACGCAGAACTGTCTACTACGGGCTGAACCTTCTTATTCCCTGTGTACTGATATCA : 681
cTGc ga TA gA aT T c T T g cG C cT T cTac T aaccT aTcaT CC TG T aT tc
halpha-6 : TTTCTAACCGTGTTGGTCTTTTACCTTCCTTCGGCTGTGGTGTGAAAAAGTGACGCTTTGTATTTCAGTCCTGCTTTCTCTGACTGTGTTTTTGCTGGTCATCA : 783
halpha-3 : TTGCTCACTGTGGCTCGTCTTCTACGTGCCCTCCCGACTGCGGTGAGAAAGGTGACCCTGTGCATTTCTGTCCTCCTCTCCCTGACGGTGTTTCTCCTGGTGATCA : 783
halpha-4 : TGCCCTCACCGTGCTGGTCTTCTACGTGCCCTCCCGAGTGTGGCGAGAAGATCACGCTGTGTGCATCTCCGTGCTGCTGTGTCGCTCACGGTCTTCCTGCTGCTCATCA : 786
halpha-2 : TGCCCTCACTGTGCTGGTCTTCTACGTGCCCTCCCGAGTGCGGCGAGAAGATCACGCTGTGTGCATTTCGGTGCTGCTGTGTCACTCACGGTCTTCCTGCTGCTCATCA : 783
halpha-5 : TTTTTAACTGTACTTGTCTTCTATCTTCCTTCAAATGAAGGTGAAAAGATTTGTGTCTGGACTTCAGTACTTGTGTCTTTGACTGTCTTCCTTCTGGTTATTG : 766
halpha-1 : TTCTTAACTGGCCTGGTAATCTAGCTGCCCACAGAGTCAGGGGAGAAGATGACTCTGAGGATCTCTGTCTTACTGTCTTTGACTGTGTTCCTTCTGGTCATCG : 786
msalpha-1 : TTCCTGACGGTCCTGGTAATTTACTTGCCTTCGGCAGTGGCGAGAAGGTATCGCTGTCCATATCGATCCTACTGTCGCTCACGCTGTTCTTCCTGCTGTTGG : 815
Dmalpha-3 : TTTCTCACCATCCTCGTAATTCTACGTGCCATCGGATAGTGGCGAGAAGGTCTCATTAAGCATATCCATTCTGCTGTCGCTCACTGTGTTCTTCCTGCTGCTGG : 811
mpalpha-1 : TACCTGTCGGTGCTGGTCTTCTATCTGCCTGCCGAGTCCAAGGAGAAAATTTCACTGTGGCATCACGATCCTCTTGTCCCAGACGCATGTTCTTCTTGCTCATCT : 810
sgalpha-1 : TACCTCTCGGTGCTGGTGTTCTACGTGCCGGCCGAGTCGGGGCGAGAAGATCGCGCTCTGGCATCTCGATCCTGGTGTGCGCAGACGCATGTTCTTCCTGCTCATAT : 813
dmalpha-2 : TATCTTTCGGTGCTGGTCTTTTACGTGCCCGCCGAATTGTGGCGAAAAGATTGCTCTGTGTGGCATCAGCATCCTGGTGTGCGCAAACGCATGTTCTTCCTGCTGCTGG : 808
mp-alpha-2 : TTCCTGTCCGTGCTGGTGTTCTACGTGCCATCCGAGTCCGGCGAGAAGGTGTCGCTGTGGCATCAGCATCCTGCTCTGCGCTTACGCTGTTCTTCCTCCTGCTGG : 811
dmalpha-1 : TTCCTGTCCGTTGTGGTCTTCTACGTGCCCAGCGAGTCTGGCGAGAAGATCTCGCTGTGGCATCAGCATCCTGCTCTGCGCTTACGCTGTTCTTCCTCCTGCTGG : 811
SEQID1 : TCCATGGCCTTGCTCGGATTCACCGTGCCCGCCAGATTCGGGGTGAAAAATTATCGCTGGGTGTTACCATCTTGGTCTGCGCTGACCGTGTTTCTGAATATGGTTG : 782
SEQID3 : TCCATGGCTCTATTGGGGTTCACCTTGCCTCCAGACTCCGGAGAAAAGTTGTCTTTAGGTGTGACGATATTACTGTCGTGTGACGGTGTTCCTCAACATGGTGG : 774
SEQID5 : TCGATGGCACTCCTCGGCTTCACACTGCCACCAGAGTCCGGAGAGAAACTCACACGTTGGAGTCACTATTCTTCTGATCGCTGACGGTGTTCCTCAACCTGGTAG : 780
halpha-7 : GCCCTCGCCCTGCTGGTGTTCCTGCTTCCTGCAGATTCCGGGGAGAAGATTTCCCTGGGGATAACAGTCTTACTCTCTCTTACGGTCTTCATGCTGCTCGTGG : 784
chalpha-7 : GCACTTGCCTTATTAGTCTTTCTGCTTCCAGCAGACTCAGGAGAAAAGATCTCACTAGGTATAACAGTTTTATTGTCTCTCACGGTCTTCATGTTACTCGTGG : 784
t T C t cT Gt TTc cTgCC c gA t gg GA AAg T c cT g at c T T cT TC ct AC gT TTc T t T T


840
860
880
900
920
halpha-6 : 886
halpha-3 : 886
halpha-4 : 889
halpha-2 : 886
halpha-5 : 869
halpha-1 : 877
msalpha-1 : 918
Dmalpha-3 : 914
mpalpha-1 : 912
sgalpha-1 : 913
dmalpha-2 : 916
mp-alpha-2 : 911
dmalpha-1 : 914
SEQID1 : 884
SEQID3 : 877
SEQID5 : 883
halpha-7 : 887
chalpha-7 : 887
940
960
980
1000
1020
halpha-6 : 893
halpha-3 : 902
halpha-4 : 905
halpha-2 : 901
halpha-5 : 882
halpha-1 : -
msalpha-1 : 934
Dmalpha-3 : 925
mpalpha-1 : 928
sgalpha-1 : 929
dmalpha-2 : 932
mp-alpha-2 : 927
dmalpha-1 : 930
SEQID1 : 901
SEQID3 : 887
SEQID5 : 980
halpha-7 : 903
chalpha-7 : 990

#28
E

ATTACHMENT 3

Molecular Cloning

A LABORATORY MANUAL

SECOND EDITION

Printed in the United States of America

9 8 7 6 5 4 3 2 1

Book and cover design by Emily Harste

Cover: The electron micrograph of bacteriophage λ particles stained with uranyl acetate was digitized and assigned false color by computer. (Thomas R. Broker, Louise T. Chow, and James L. Garrels)

Cataloging in Publications data

Sambrook, Joseph
Molecular cloning : a laboratory manual / E.F. Fritsch, T. Maniatis—2nd ed.
p. cm.
Bibliography: p.
Includes index.
ISBN 0-87969-309-6
1. Molecular cloning—Laboratory manuals. 2. Eukaryotic cells-Laboratory manuals. I. Fritsch, Edward F. II. Maniatis, Thomas III. Title.
QH442.2.M26 1987
574.87'3224—dc19 87-35464

All Cold Spring Harbor Laboratory Press publications may be ordered directly from Cold Spring Harbor Laboratory, Box 100, Cold Spring Harbor, New York 11724. Phone: 1-800-843-4388. In New York (516)367-8423.

IDENTIFICATION OF cDNA CLONES OF INTEREST

Methods of Screening

There are three methods to screen cDNA libraries for clones of interest:

- Nucleic acid hybridization
- Immunological detection of specific antigens
- Sib selection either by hybrid selection and translation of mRNA or by production of biologically active molecules

Most cloning projects today are aimed at isolating cDNAs corresponding to rare mRNAs and therefore require screening of large numbers of recombinant clones. This can be carried out effectively with only two types of reagents: antibodies and nucleic acid probes. In those rare instances when both types of reagents are available, nucleic acid probes are preferred because they can be used under a variety of different stringencies that minimize the chance of undesirable cross-reactions. Furthermore, nucleic acid probes will detect all clones that contain cDNA sequences, whereas antibodies will react only with a subset of these clones (in some cases one in six at best) in which the cDNA has been inserted into the vector in the correct reading frame and orientation. cDNA libraries that are to be screened by antibodies therefore need to be larger (by a factor of at least 6) than those that are to be screened by nucleic acid probes. Consequently, when using antibody probes to search for a cDNA clone corresponding to a mammalian mRNA present at the level of 1 molecule/cell or less, it is desirable to construct cDNA expression libraries that contain in excess of 10^7 members. This is not easy, especially when the amounts of mRNA are limited. Furthermore, screening a library of this size is expensive and laborious, and it becomes worthwhile to explore methods to enrich the mRNA (or cDNA derived from it) for the sequences of interest (see pages 8.6–8.10).

NUCLEIC ACID HYBRIDIZATION

This is the most commonly used and reliable method of screening cDNA libraries for clones of interest. None of the other methods displays such an abundance of attractive features. Screening by nucleic acid hybridization allows extremely large numbers of clones to be analyzed simultaneously and rapidly, does not require that the cDNA clones be full-length, and does not require that an antigenically or biologically active product be synthesized in the host cell. Furthermore, as a result of more than 20 years of work, the theoretical basis of nucleic acid hybridization is well-understood. This has led to the development of a large number of different techniques that can accommodate nucleic acid probes of very different lengths and specificities. Details of the methods for the preparation and use of these probes are presented in Chapters 10 and 11.

Homologous probes

Homologous probes contain at least part of the exact nucleic acid sequence of the desired cDNA clone. They are used in a variety of circumstances, for

example, when a partial clone of an existing cDNA is used to isolate a full-length clone from a cDNA library. Usually, a fragment derived from one nd or the other of the existing clone is isolated, radiolabeled in vitro, and used to probe a library. Hybridization with homologous probes is always carried out under stringent conditions.

Partially homologous probes

Partially homologous probes are used to detect cDNA clones that are related, but not identical, to the probe sequences. If neither antibody nor nucleic acid probes are available, a number of alternative strategies can be considered. For example, if the same gene has already been cloned from another species or if a related gene has been cloned from the same species, it would be worthwhile carrying out a series of trial experiments to determine whether there is sufficient conservation of nucleic acid sequence to allow the screening of a cDNA library by hybridization. This is most easily accomplished by performing a series of Southern and northern hybridizations at different stringencies. For example, a large batch (50 μg) of genomic DNA is cleaved with a restriction enzyme that cleaves the probe sequence at one or two well-separated sites. It is a good idea to digest an equal amount of genomic DNA of the original species for use as a positive control. Aliquots (5–10 μg) of the digests are then applied to adjacent slots of a 0.8% agarose gel, electrophoresis is carried out, and the fragments are then transferred to a nitrocellulose filter as described in Chapter 9, pages 9.34–9.41. The filter is cut into strips, each of which is hybridized under different conditions to identical amounts of radioactive probe. For aqueous hybridization, the ionic strength of the solution is kept constant (usually 1 M Na^+) while the temperature of annealing is progressively lowered (from 68°C to 42°C). The strips are then washed extensively at the temperature of hybridization with a solution containing 2 × SSC, 0.5% SDS. When hybridization is carried out in solvents containing formamide, the temperature and ionic strength are usually kept constant (42°C and 6 × SSC [or 6 × SSPE], respectively) while the amount of formamide in the annealing buffer is progressively lowered from 50% to 0%. The strips are then washed extensively at 50°C in 6 × SSC, 0.5% SDS. A similar series of hybridizations can be carried out with mRNA preparations that have been fractionated by electrophoresis and transferred to a solid support. In both cases, the aim is to establish conditions that will allow the previously cloned gene to be used as a probe for the cDNA of interest, without undue interference from background hybridization.

Total cDNA probes

Total cDNA probes are prepared by uniform incorporation of radiolabeled nucleotides with reverse transcriptase or end-labeling of total or fractionated poly(A)$^+$ mRNA. They can be used to screen libraries of cDNA for specific clones *if* the cDNA clones of interest correspond to mRNA species present in the initial population at a frequency of at least 1 in 200 (see Gergen et al. 1979; Dworkin and Dawid 1980). It is not possible to detect cDNA clones homologous to species that are represented rarely in the mRNA preparation.

HYBRIDIZATION OF RADIOLABELED PROBES TO IMMOBILIZED NUCLEIC ACIDS

There are many methods available to hybridize radioactive probes in solution to nucleic acids immobilized on solid supports such as nitrocellulose filters or nylon membranes. These methods differ in the following respects:

- Solvent and temperature used (e.g., 68°C in aqueous solution or 42°C in 50% formamide)
- Volume of solvent and length of hybridization (large volumes for periods as long as 3 days or minimal volumes for periods as short as 4 hours)
- Degree and method of agitation (continuous shaking or stationary)
- Use of agents such as Denhardt's reagent or BLOTTO to block the non-specific attachment of the probe to the surface of the solid matrix
- Concentration of the labeled probe and its specific activity
- Use of compounds, such as dextran sulfate (Wahl et al. 1979) or polyethylene glycol (Renz and Kurz 1984; Amasino 1986), that increase the rate of reassociation of nucleic acids
- Stringency of washing following the hybridization

Although the choice depends to a large extent on personal preference, we offer the following guidelines for choosing among the various methods available.

1. Hybridization reactions in 50% formamide at 42°C are less harsh on nitrocellulose filters than is hybridization at 68°C in aqueous solution. However, it has been found that the kinetics of hybridization in 80% formamide are approximately four times slower than in aqueous solution (Casey and Davidson 1977). Assuming a linear relationship between the rate of hybridization and the formamide concentration, the rate in 50% formamide should be two to three times slower than the rate in aqueous solution. Both types of solvents give excellent results and neither has a clear-cut advantage over the other.
2. The smaller the volume of hybridization solution, the better. In small volumes of solution, the kinetics of nucleic acid reassociation are faster and the amount of probe needed can be reduced so that the DNA on the filter acts as the driver for the reaction. However, it is essential that sufficient liquid be present for the filters to remain covered at all times by a film of the hybridization solution.
3. Continual movement of the probe solution across the filter is unnecessary, even for a reaction driven by the DNA immobilized on the filter. However, if a large number of filters are hybridized simultaneously, agitation is advisable to prevent the filters from adhering to one another.
4. The kinetics of the hybridization reaction are difficult to predict from theoretical considerations, partly because the exact concentration of the

immobilized nucleic acid and its availability for hybridization are unknown. When using probes that have the capacity to self-anneal (e.g., nick-translated double-stranded DNA), the following rule of thumb is useful: Allow the hybridization to proceed for a time sufficient to enable the probe in solution to achieve $1-3 \times C_o t_{1/2}$. In 10 ml of hybridization solution, 1 μg of a probe of 5-kb complexity will reach $C_o t_{1/2}$ in 2 hours. To determine the time of half-renaturation for any other probe, simply enter the appropriate values into the following equation:

$$1/x \times y/5 \times z/10 \times 2 = \text{number of hours to achieve } C_o t_{1/2}$$

where x = the weight of the probe added (in micrograms), y = its complexity (for most probes, complexity is proportional to the length of the probe in kilobases), and z = the volume of the reaction (in milliliters).

After hybridization to $3 \times C_o t_{1/2}$ has been reached, the amount of probe available for additional hybridization to the filter is negligible. For probes that do not have the capacity to self-anneal (e.g., RNA probes synthesized in vitro by bacteriophage-encoded DNA-dependent RNA polymerases; see Chapter 10), the hybridization time may be shortened, since the lack of a competing reaction in the solution favors hybridization of the probe to the DNA immobilized on the filter.

5. Several different types of agents can be used to block the nonspecific attachment of the probe to the surface of the filter. These include Denhardt's reagent (Denhardt 1966), heparin (Singh and Jones 1984), and nonfat dried milk (Johnson et al. 1984). Frequently, these agents are used in combination with denatured, fragmented salmon sperm or yeast DNA and detergents such as SDS. In our experience, virtually complete suppression of background hybridization is obtained by prehybridizing filters with a blocking agent consisting of 5× Denhardt's reagent, 0.5% SDS, and 100 μg/ml denatured, fragmented DNA. We recommend this mixture whenever the signal-to-noise ratio is expected to be low, for example, when carrying out northern analysis of low-abundance mRNAs or Southern hybridizations with single-copy sequences of mammalian DNA. However, in most other circumstances (Grunstein/Hogness hybridization [1975], Benton/Davis hybridization [1977], Southern hybridization [1975] of abundant DNA sequences, etc.), we recommend using 0.25% nonfat dried milk (0.05× BLOTTO; Johnson et al. 1984). This is much less expensive, easier to use than Denhardt's reagent, and, for these purposes, gives results that are equally satisfactory. In general, Denhardt's reagent is more effective for nylon membranes. The signal-to-noise ratio obtained with most brands of nylon membranes is higher with Denhardt's reagent than with BLOTTO. Nonfat dried milk is not recommended when using RNA probes or when carrying out northern hybridizations because of the possibility that it might contain unacceptably high levels of RNAase activity. For more information about blocking agents, see Table 9.1.

6. Blocking agents are usually included in both the prehybridization and hybridization solutions when nitrocellulose filters are used. However, when the nucleic acid is immobilized on nylon membranes, the blocking agent- are often omitted from the hybridization solution, since high

concentrations of protein are believed to interfere with the annealing of the probe to its target. This quenching of the hybridization signal is particularly noticeable when oligonucleotides or probes less than 100 nucleotides in length are used.

7. In the presence of 10% dextran sulfate or 10% polyethylene glycol, the rate of hybridization is accelerated approximately tenfold (Wahl et al. 1979; Renz and Kurz 1984; Amasino 1986) because nucleic acids are excluded from the volume of the solution occupied by the polymer and their effective concentration is therefore increased. Although dextran sulfate and polyethylene glycol are useful in circumstances where the rate of hybridization is the limiting factor in detecting rare sequences (e.g., northern or genomic Southern blots), they are of no benefit when screening bacterial colonies or bacterial plaques. In addition, they can sometimes lead to high backgrounds, and hybridization solutions containing them are always difficult to handle because of their viscosity. We therefore recommend that dextran sulfate and polyethylene glycol not be used unless the rate of hybridization is very slow, the filter contains very small amounts of DNA, or the amount of radiolabeled probe is limiting.

8. To maximize the rate of annealing of the probe with its target, hybridizations are usually carried out in solutions of high ionic strength (6 × SSC or 6 × SSPE) at a temperature that is 20–25°C below the melting temperature (T_m). Both solutions work equally well when hybridization is carried out in aqueous solvents. However, when formamide is included in the hybridization buffer, 6 × SSPE is preferred because of its greater buffering power.

9. In general, the washing conditions should be as stringent as possible (i.e., a combination of temperature and salt concentration should be chosen that is approximately 12–20°C below the calculated T_m of the hybrid under study). The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the probe of interest and then washed under conditions of different stringencies.

10. To minimize background problems, it is best to hybridize for the shortest possible time using the minimum amount of probe. For Southern hybridization of mammalian genomic DNA where each lane of the gel contains 10 μg of DNA, 10–20 ng/ml radiolabeled probe (sp. act. = 10^9 cpm/μg or greater) should be used and hybridization should be carried out for 12–16 hours at 68°C in aqueous solution or for 24 hours at 42°C in 50% formamide. For Southern hybridization of fragments of cloned DNA where each band of the restriction digest contains 10 ng of DNA or more, much less probe is required. Typically, hybridization is carried out for 6–8 hours using 1–2 ng/ml radiolabeled probe (sp. act. = 10^9 cpm/μg or greater).

11. *Useful facts*:

    a. The T_m of the hybrid formed between the probe and its target may be estimated from the following equation (~~Bolton and McCarthy 1962~~ Gleichung eigentl.: Bonner et al '73).

$$T_m = 81.5°C - 16.6(\log_{10}[Na^+]) + 0.41(\%\,G + C) - 0.63(\%\,\text{formamide}) - (600/l)$$

where l = the length of the hybrid in base pairs.

This equation is valid for:

- Concentrations of Na^+ in the range of 0.01 M to 0.4 M. It predicts T_m less accurately in solutions of higher $[Na^+]$.
- DNAs whose G + C content is in the range of 30% to 75%. Note that the depression of T_m in solutions containing formamide is greater for poly(dA:dT) (0.75°C/1% formamide) and less for DNAs rich in poly(dG:dC) (0.50°C/1% formamide) (Casey and Davidson 1977).

The equation applies to the "reversible" T_m that is defined by optical measurement of hyperchromicity at OD_{267}. The "irreversible" T_m, which is more important for autoradiographic detection of DNA hybrids, is usually 7–10°C higher than that predicted by the equation.

Similar equations have been derived for:

i. RNA probes hybridizing to immobilized RNA (Bodkin and Knudson 1985)

$$T_m = 79.8°C + 18.5(\log_{10}[Na^+]) + 0.58(\%G + C) + 11.8(\%G + C)^2 - 0.35(\%\,\text{formamide}) - (820/l)$$

ii. DNA:RNA hybrids (Casey and Davidson 1977)

$$T_m = 79.8°C + 18.5(\log_{10}[Na^+]) + 0.58(\%G + C) + 11.8(\%G + C)^2 - 0.50(\%\,\text{formamide}) - (820/l)$$

Comparison of these equations shows that the relative stability of nucleic acid hybrids decreases in the following order: RNA:RNA (most stable), RNA:DNA (less stable), and DNA:DNA (least stable). In aqueous solutions, the T_m of a DNA:DNA hybrid is approximately 10°C lower than that of the equivalent RNA:RNA hybrid. In 80% formamide, the T_m of an RNA:DNA hybrid is approximately 10°C higher than that of the equivalent DNA:DNA hybrid.

b. The T_m of a double-stranded DNA decreases by 1–1.5°C with every 1% decrease in homology (Bonner et al. 1973).

The above equations apply only to hybrids greater than 100 nucleotides in length. The behavior of oligonucleotide probes is described in detail in Chapter 11.

For a general discussion of hybridization of nucleic acids bound to solid supports, see Meinkoth and Wahl (1984).

References

Amasino, R.M. 1986. Acceleration of nucleic acid hybridization rate by polyethylene glycol. *Anal. Biochem.* **152:** 304.

Arnheim, N. and E.M. Southern. 1977. Heterogeneity of the ribosomal genes in mice and men. *Cell* **11:** 363.

Benton, W.D. and R.W. Davis. 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. *Science* **196:** 180.

Bittner, M., P. Kupferer, and C.F. Morris. 1980. Electrophoretic transfer of proteins and nucleic acids from slab gels to diazobenzyloxymethyl cellulose or nitrocellulose sheets. *Anal. Biochem.* **102:** 459.

Blin, N. and D.W. Stafford. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Res.* **8:** 2303.

Bodkin, D.K. and D.L. Knudson. 1985. Assessment of sequence relatedness of double-stranded RNA genes by RNA-RNA blot hybridization. *J. Virol. Methods* **10:** 45.

Bolton, E.T. and B.J. McCarthy. 1962. A general method for the isolation of RNA complementary to DNA. *Proc. Natl. Acad. Sci.* **48:** 1390.

Bonner, T.I., D.J. Brenner, B.R. Neufeld, and R.J. Britten. 1973. Reduction in the rate of DNA reassociation by sequence divergence. *J. Mol. Biol.* **81:** 123.

Bos, J.L., M. Verlaan-de Vries, A.M. Jansen, G.H. Veeneman, J.H. van Boom, and A.J. van der Eb. 1984. Three different mutations in codon 61 of the human N-*ras* gene detected by synthetic oligonucleotide hybridization. *Nucleic Acids Res.* **12:** 9155.

Bos, J.L., E.R. Fearon, S.R. Hamilton, M. Verlaan-de Vries, J.H. van Boom, A.J. van der Eb, and B. Vogelstein. 1987. Prevalence of *ras* gene mutations in human colorectal cancers. *Nature* **327:** 293.

Bos, J.L., D. Toksoz, C.J. Marshall, M. Verlaan-de Vries, G.H. Veeneman, A.J. van d r Eb, J.H. van Boom, J.W.G. Janssen, and A.C.M. Steenvoorden. 1985. Amino-acid substitutions at codon 13 of the N-*ras* oncogene in human acute myeloid leukaemia. *Nature* **315:** 726.

Botchan, M., G. McKenna, and P.A. Sharp. 1974. Cleavage of mouse DNA by a restriction enzyme as a clue to the arrangement of genes. *Cold Spring Harbor Symp. Quant. Biol.* **38:** 383.

Bowtell, D.D.L. 1987. Rapid isolation of eukaryotic DNA. *Anal. Biochem.* **162:** 463.

Burgers, P.M.J. and K.J. Percival. 1987. Transformation of yeast spheroplasts without cell fusion. *Anal. Biochem.* **163:** 391.

Burke, D.T., G.F. Carle, and M.V. Olson. 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236:** 806.

Casey, J. and N. Davidson. 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **4:** 1539.

Church, G.M. and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci.* **81:** 1991.

Clarke, L. and J. Carbon. 1976. A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E. coli* genome. *Cell* **9:** 91.

Conner, B.J., A.A. Reyes, C. Morin, K. Itakura, R.L. Teplitz, and R.B. Wallace. 1983. Detection of sickle cell β^S-globin allele by hybridization with synthetic oligonucleotides. *Proc. Natl. Acad. Sci.* **80:** 278.

Cooke, H. 1987. Cloning in yeast: An appropriate scale for mammalian genomes. *Trends Genet.* **3:** 173.

Daniels, D.L., J.L. Schroeder, W. Szybalski, F. Sanger, A.R. Coulson, G.F. Hong, D.F. Hill, G.B. Petersen, and F.R. Blattner. 1983. Appendix II: Complete annotated lambda sequence. In *Lambda II* (ed. R.W. Hendrix et al.), p. 519. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Denhardt, D.T. 1966. A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23:** 641.